International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/223318 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

A Novel All Members Group Search Optimization Based Data Acquisition in Cloud Assisted Wireless Sensor Network for Smart Farming

Vuppala Sukanya
Department of Computer Science and Engineering, University College of Engineering, Osmania University, Hyderabad, Telangana, India.
vsukanya060@gmail.com

Ramachandram S
Department of Computer Science and Engineering, University College of Engineering, Osmania University, Hyderabad, Telangana, India.
sramachandram14@gmail.com

Received: 21 June 2023 / Revised: 10 August 2023 / Accepted: 21 August 2023 / Published: 31 August 2023

Abstract – Recent times, the Wireless Sensor Networks (WSN) has played an important role in smart farming systems. However, WSN-enabled smart farming (SF) systems need reliable communication to minimize overhead, end-to-end delay, latency etc., Hence, this work introduces a 3-tiered framework based on the integration of WSN with the edge and cloud computing platforms to acquire, process and store useful soil data from agricultural lands. Initially, the sensors are deployed randomly throughout the network region to collect information regarding different types of soil components. The sensors are clustered based on distance using the Levy flight based K-means clustering algorithm to promote efficient communication. The Tasmanian devil optimization (TDO) algorithm is used to choose the cluster heads (CHs) based on the distance among the node and edge server, residual energy, and the number of neighbors. Then, the optimal paths to transmit the data are identified using the all members group search optimization (AMGSO) algorithm based on different parameters. Each edge server assesses the quality of the data (QoD) with respect to some data quality criteria after receiving the data from the edge server. Also, the load across the servers are balanced in order to overcome the overloading and under loading issues. The legitimate data that received higher scores in the QoD evaluation alone is sent to the cloud servers for archival. Using the ICRISAT dataset, the efficiency of the proposed work is evaluated using a number of indicators. The average improvement rate attained by the proposed model in terms of energy consumption is 40%, in terms of packet delivery ratio is 7%, in terms of network lifetime is 38%, and in terms of latency is 24% for a total of 250 nodes.

Index Terms – Smart Agriculture, Wireless Sensor Network, Edge Computing, Cloud Computing, Clustering, Routing, Data Quality Evaluation, Load Balancing.

1. INTRODUCTION

Agriculture is a backbone sector of the country's economy that satisfies the basic need of human life and urges the necessity of introducing technologies to result in increased production. The cultivated lands may sometimes be affected by diverse environmental changes, and this impacts overall productivity causing losses to farmers [1]. To avoid such conditions, it is important to accurately monitor the lands and the soil contents so that further steps can be taken to increase productivity and crop quality [2]. Efficient irrigation is needed, particularly in the arid and semi-arid regions, as these regions are difficult to monitor. Only using the data gathered from the lands can the irrigation facilities in the lands be affected [3]. The irrigation procedures can be broadly categorized into three parts such as sprinkler, gravity and drip irrigation. Among these three, the gravity type is the oldest and is considered one of the poor types in conserving natural resources. It is important to accurately monitor and obtain information from the cultivated lands to choose between these irrigation types [4, 5].

Precision agriculture is a procedure where smart decisions can be taken based on some monitored data to improve overall productivity [6]. This concept is dependent on the management of crops through observation, measurement and acting based on the changes in certain parameters that impact the growth of crops [7]. One of the recent innovations in the field of smart farming is the integration of wireless sensor networks (WSNs) that effectively monitor lands with the placement of sensors. To sense the necessary characteristics, such as soil nutrients or other air components, the sensors are deployed directly on agricultural areas [8, 9]. The base station receives the detected data and processes it further. The data collected from the fields are then provided to some decision-making systems to take appropriate decisions regarding the irrigation type required or the requirement of water or other resources for the betterment of crops [10, 11]. Monitoring crop lands becomes essential when the soil or other environmental parameters must be adjusted or controlled to reduce the negative impacts. If the deployed system can closely monitor the data from the fields, the overall performance can be improved [12].

Energy consumption concerns that shorten the network's life and cause data transmission failures are one of the key issues with precision farming. The battery life of the sensor nodes is limited, and each communication uses energy, making battery drainage simple [13]. To avoid this, the preferred paths for transferring the data to the base station can be optimized. When the chosen paths are affected by high traffic load due to the dynamic nature of the network, more energy is required to be spent to transmit the data [14]. At this moment, the network must determine the optimal paths that are close to the base station with the fewest hops so that the least amount of energy is expended for data transmission [15]. To identify the optimal paths, several methodologies are formulated with different objectives. A move based on metaheuristic optimization is made here to identify the optimal paths for transmitting the sensor collected data. The proposed method combines two effective metaheuristics to achieve the desired objective.

1.1. Motivation

Though there are several works developed with the motive of attaining efficient data acquisition in smart agriculture, there is a lack of proficient techniques that can attain performance enhancement in terms of several aspects. Most of the algorithms introduced in the literature just focused on collecting and forwarding the data to the servers or base stations (BS) which is only a part of the concern in this field. Certain techniques result in efficient data acquisition results. Still, those techniques cannot determine the attendance of dead nodes in the network, which may abruptly lead to a reduced network lifetime. Besides this, continuous loading of edge servers may result in overloading or under loading issues considered in most of the established techniques. All these research gaps motivated to propose an optimized framework for data acquisition to attain efficient smart agriculture.

1.2. Contribution

The major contributions of this work are highlighted below:

- Design a framework integrating cloud computing technology with the WSN network to enable effective precision farming.
- Formulate some influential parameters such as execution cost, resource utilization, latency, response time, and energy consumption into objective functions to compute the suitable path for data transmission.
- The most optimal nodes in the network are selected as CHs with respect to residual energy, number of neighbors, and distance using the Tasmanian Devil Optimization (TDO) algorithm.
- Presented a novel metaheuristics-based optimization framework called the Hybrid All Members Group Search Optimization (AMGSO) model to find the optimal paths for transmission of the collected agricultural field data.
- The proposed framework optimally identifies the dead nodes in the path and finds an alternative path in every iteration to avoid disruption in communication.
- The load on the edge servers is balanced based on the resource availability on the edge servers to avoid overload and underload issues.
- Extensive simulation and assessment of the proposed work using the ICRISAT dataset based on various metrics to demonstrate its performance effectiveness compared to other data collection models.

1.3. Paper plan

The structure of the work is as follows: Section 2 provides the past studies explored under this topic, Section 3 explains the proposed work with algorithms in detail, the outcomes and discussions is provided in Section 4, and the conclusion of the work is presented in Section 5.

2. RELATED WORK

In literature, various studies have been carried to improve the future of precision farming, and some of the effective and recent ones are reviewed below:

Monitoring crops grown in agricultural fields is crucial as it helps increase the overall production rate. When the area considered is huge in size, it is difficult to exactly monitor the crops, leading to some technological requirements. The recent integration of WSN technology with precision farming helped gather the required field information much more effectively. The sensors are positioned in the fields to collect the required data for further processing. Based on this idea, Khelifi [16] introduced a methodology where atmospheric sensors were deployed in the fields to gather the air and soil components. A periodic hybrid routing algorithm (PHRA) was introduced for collecting the data, which was sensitive to the threshold. In order to deploy the sensors on the fields that ensured effective coverage of the considered agricultural area, the procedure of regional clustering was followed. For the purpose of electing the cluster head (CH), a clustering protocol was formulated in terms of parameters such as distance and residual energy. The performance of the method was proven by simulations, and it was found that the method stores energy in the network over a longer period of time.

The energy loss in sensor nodes during data transmission and communication is one of the primary issues in integrating WSNs for precision farming. This is owing to the sensor nodes' restricted capacity of battery and the energy used for each transmission, which causes a power drain. This problem can be resolved with an optimal network configuration that supports effective communication without any expenditure of energy. A methodology to achieve optimal network configuration for the cultivation of ginseng was introduced by Li et al. [17]. A metaheuristic based networking algorithm was introduced to maintain optimal network configuration with the balance of energy and effective communication. The particle swarm optimization (PSO) algorithm was utilized to achieve the objective. The routes were created, and the optimal routes were identified using the PSO algorithm, resulting in the minimum energy expenditure throughout the network. Experimental evaluations proved that the methodology was superior to most existing routing-based methodologies.

Several parameters must be considered when searching for the optimal paths for data transmission. One of the main problems in data transmission is the dynamics of the network due to traffic changes. When the network load is changed, the adaptability of the network to the load should be noted. With changes in network traffic, the selected paths for data transmission get loaded, thereby degrading the performance. To avoid this, appropriate solutions are required to be formulated. A suitable solution was implemented by Agarkhed et al. [18] to enable effective communication in the network for precision farming. The method was based on intelligent traffic path selection so that the data could be transmitted only through optimal paths without the degradation of performance. The network nodes were clustered, and the CH was chosen based on the highest residual energy maintained by the nodes. Then, based on the network traffic rate, a decision support system (DSS) was deployed to support reliable data transfer. The system was implemented and assessed using several methodologies, and the results demonstrated that the system was optimal.

The major requirement with the integration of WSN and precision farming is the improvement of network lifespan by reducing the over consumption of energy. To resolve the issue, an intelligent routing system was put forth by Pandiyaraju et al. [19] based on fuzzy rules called the terrain based routing protocol. Initially, terrains were formed by partitioning the far field into several equal sized lands with the assurance of effective coverage. The sensors were randomly deployed in the terrains, and a terrain head (TH) was selected for every terrain based on the distance and remaining energy. Based on the node's distance from the sink and the present energy level, fuzzy logic was used to select the TH. Finally, the routing phase was carried out either in a single-hop or multi-hop fashion and the data was transmitted through the selected route with the help of relay nodes. Again the fuzzy logic was applied to select the next hop relay nodes to maintain energy in the network. The experimental outcomes revealed that the model was more effective and reliable in increasing irrigation than other existing work.

Providing environmentally friendly farming methodologies is crucial to boost irrigation and cultivation in the modern world. With the deployment of sensor nodes in agricultural lands, the data collection and transmission facilities have improved a lot. But some complex problems, such as energy issues due to limited battery power and complex routing scenarios resulting in delayed and failed transmissions, require attention. These problems result in extra burden in base stations that require prompt solutions. One approach was designed by Qureshi et al. [20] to provide a reliable solution to the issue above. The method developed the gateway clustering based energy aware centroid routing protocol that effectively forwarded the data gathered to the base station. The protocol selected the CH based on the centroid location within the cluster, and the gateway nodes were selected from each cluster. The selected gateway nodes were responsible for forwarding the collected data to the base station by reducing the data load of CH nodes. The simulations proved that the approach significantly outperformed several existing methodologies and ensured feasible monitoring of cultivated lands.

An energy efficient protocol for precision Agriculture was developed by Yao et al. [21] based on concept of multi-threshold segmentation. In cluster formation phase, a new node clustering technique was executed based on the inspiration of multi-threshold image segmentation. In cluster head selection phase, the optimal count and position of cluster head was calculated by employing new cluster head selection techniques with respect to sensor node energy and position. The developed protocol was reduced the network energy utilization and improved the quality of the cluster creation and cluster head selection.

Mahajan et al. [22] presented the industry 4.0 supported light weight clustering protocol for modern agriculture system. In the stage of clustering, the primary and secondary parameters of sensor node was computed to create the combined value of fitness function. The combined fitness function was computed for increasing network performance by minimizing energy and overhead. Especially, the Bacterial Foraging Algorithm (FBA) was used to select the optimal cluster head based on combined fitness function.

Kumar et al. [23] presented a cross layer based stable routing protocol for IoT-based smart farming environment. In

RESEARCH ARTICLE

clustering phase, cross layer parameters related with network and access control were utilized to calculate the applicability of every sensor node. The hybrid particle Swarm wild horse optimizer (PSWHO) was employed for optimal cluster head selection. Moreover, the golden eagle optimization (GEO) based deep neural network (DNN) was developed to discover the ideal route for transferring the data. Table 1 presents a comparative examination of the available literature works.

Table 1 Comparative Analysis of the Past Literature Studies

References	Methods	Purpose	Advantages	Disadvantages
Khelifi [16]	Region based clustering, periodic hybrid routing	Energy efficient data transmission	Achieved higher coverage rate along with energy efficiency	The method is based on the TDMA protocol, where time intervals are assigned for transmission. This interrupts communication as the time slots may sometimes get filled.
Li et al. [17]	PSO	Energy efficient and optimal data transmission	The labour costs are reduced with the implementation of this model, even in larger fields.	Only a few components are considered, which is insufficient to raise the productivity of the crop.
Agarkhed et al. [18]	DSS	Optimal data transmission	Intelligent path selection scheme is followed that reduces the delay in transmission by increasing the path count based on the data arrival rate.	Rule-based algorithm is followed, which is highly time-consuming for larger, more challenging agricultural lands.
Pandiyaraju et al. [19]	Fuzzy rules, terrain based routing	Energy efficient data transmission	The method conserves more energy, thereby enhancing the network lifetime.	The major drawback of this system is that the rules must be updated frequently for changing environments which is highly time-consuming.
Qureshi et al. [20]	GCEEC	Energy efficient data transmission	The method reduced the overall load in BS by administering gateway nodes.	Rotation overhead occurs when the distance between CHs and BS increases.
Yao et al. [21]	Multi-threshold segmentation	Effective cluster formation and cluster head selection to sustain the energy efficiency	Reduce energy efficiency and prolong network lifetime	Failed to achieve reliable data transmission

RESEARCH ARTICLE

Mahajan et al.[22]	Bacterial Foraging Algorithm (FBA)	Energy efficient clustering and routing	Minimize energy consumption and network overhead	Failed to ensure security during data transmission
Kumar et al.[23]	PSWHO and GEO based DNN	Energy and QoS efficient data transmission	Minimized energy consumption, delay and latency	Not suitable for load balancing during heavy traffic

Table 1 provides the comparative analysis of clustering and routing techniques of past studies explored in WSN based precision farming. Initially, the region based clustering and periodic routing technique was used in [16], which delivers good outcomes in terms of energy consumption and delivery ratio; but, it fails to select optimal cluster head. PSO algorithm was explored in [17] that provide less energy utilization, minimal delay and fair network lifespan; but it holds poor convergence speed and exploitation ability. Furthermore, an intelligent path selection scheme was employed in [18] that offer higher throughput and packet delivery ratio. However, it is fails to support dynamic cluster head selection. Fuzzy rule based routing method is presented in [19] that provides better network lifetime. Though, it needs more consideration on network latency. Finally, the optimized cluster and routing based technique was explored in [20-23], which has the good enhancement in a lifespan and minimize energy consumption for while data forwarding and receiving, but fails to provide multi-hop communication. Hence, to conquer the drawbacks of existing studies, an enhanced cluster head and route selection method is developed in this study based on the combination of bio and nature inspired metaheuristic algorithms.

2.1. Problem Statement

After examining the existing works, it was found that they cannot offer a complete increase in performance for smart farming. Though certain techniques integrated the computing platforms for better and more effective results, the problems encountered through this integration are not well resolved. One such problem is the overloading or under loading issue that may arrive with the continuous transmission of data to the edge servers. Another problem is that the WSN may involve dead nodes that disrupt the communication as the dead nodes cannot transmit the data packets further. Though the dead node is an old problem, this is only overlooked in most existing data acquisition frameworks. Therefore, this work presents an optimization based framework to effectively resolve the existing issues and to attain efficient data acquisition for smart agriculture.

3. PROPOSED METHODOLOGY

The proposed method presents an optimized method for data collection to support precision farming. In this work, the WSN, an edge and cloud-based precision farming system, is considered in which the sensor nodes are randomly placed in the agricultural areas. The sensor nodes on the fields collect the soil data and transmit it to the destination for further processing. The nodes in the fields are considered clustered to make the network more efficient. After cluster formation, optimal CHs are selected for each cluster using the TDO algorithm. Then the CHs within each cluster transfer the data to the edge servers. The edge servers evaluate data quality (QoD) and then forward the data to the cloud system for storage. The architecture of the proposed work is shown in Figure 1.

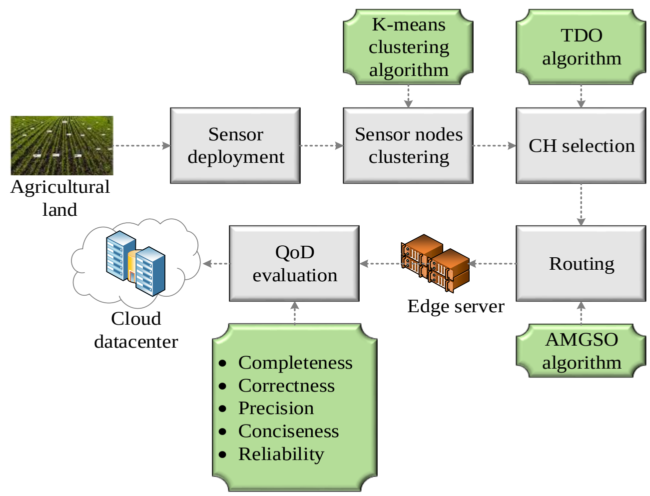


Figure 1 Architecture of the Proposed Work

The nodes placed in the lands are clustered in the suggested framework using the k-means clustering algorithm. Based on the distance among the cluster and the edge server, the centroid is selected as the CH for each cluster. After clustering the nodes, the sensors in each cluster collect the data and forward it to the respective CHs for transmission. The CH in each cluster identifies the optimal path to the edge server using the hybrid AMGSO algorithm based on certain objectives. When searching for the optimal path, the dead nodes present in the monitoring environment are detected based on their energy level, the respective path is discarded, and the algorithm identifies an alternative path. After reaching the edge server, the data is distributed to the available servers to reduce the overloading and under loading issues. Only the legitimate data is sent to the cloud system for storage and

RESEARCH ARTICLE

analysis after the edge server evaluates the QoD of the data using various quality indicators.

3.1. System Model

The system model followed in this work is three-tiered and integrates the WSN into the edge and cloud computing platforms. The first tier involves the WSN model, which deploys the sensor nodes on agricultural land to collect soil data. These sensors are placed randomly, and all sensors are assumed to have the same detection capability. The sensors considered in this work are of different types, and each type is intended to detect a specific soil component. The second tier includes the edge servers, which are deployed for remote data collection from sensor nodes. The third tier is the cloud tier, where the servers reside, and all the collected and valid data is forwarded to the servers for storage. The system model pursued in the work is shown in Figure 2.

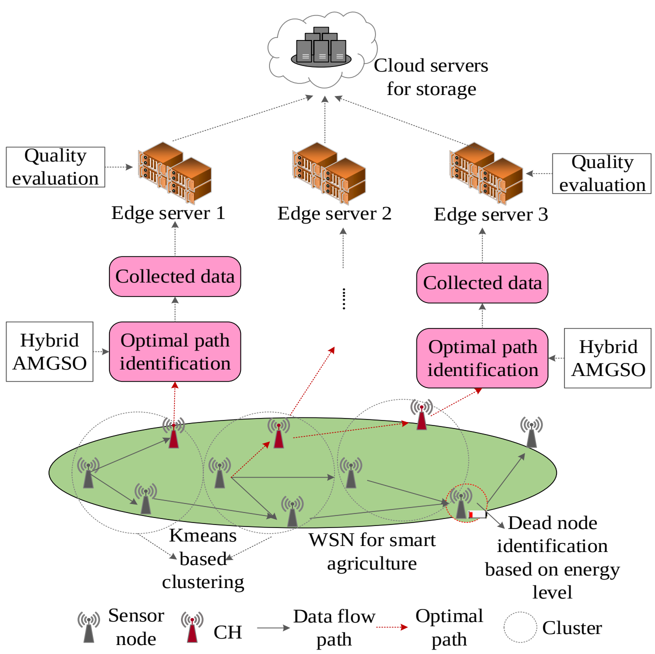


Figure 2 System Model of the Proposed Work

The network's nodes are grouped according to their distance from one another, and the CHs then gather the data and send it from the cluster to the edge server. The data is transferred to the edge server via optimal paths identified by the routing algorithm. Then the QoD of the data is evaluated, and only the valid data is retained and transmitted to the cloud, while the invalid data is discarded.

3.2. Clustering

All nodes deployed in the network are clustered based on their distance in the first step of the proposed work. To group the nodes and enable efficient data transmission, the proposed work follows the levy flight based k-means clustering algorithm which is more optimal and effective in determining the clusters than traditional clustering algorithms. The process followed in levy flight based k-means is as follows:

3.2.1 Levy Flight Based K-Means for Clustering the Sensor Nodes

The sensor nodes are clustered in the proposed work using the Levy flight based k-means clustering algorithm. This algorithm has the prime advantages of better convergence and generalization even on larger datasets. Moreover, this algorithm is highly adaptable and can be implemented to solve any clustering problem. The proposed work follows this algorithm to cluster the randomly deployed sensor nodes in the field. Especially, the Levy flight behaviour is used to find the new positions to evade premature convergence in clustering. The steps followed in the construction of clusters are as follows:

Step 1: At the initial stage, the k-number of centroids is chosen randomly from different places in the solution space.

Step 2: Determine the distance between the individual nodes in the solution space and also find the centroids. Based on the computed distance, the nodes are assigned to the nearest centroid, which results in the formation of 'k' number of clusters. To determine the distance between the nodes, the Euclidean distance measure is computation based on equation (1).

$$D = \sqrt{\left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right]} \quad (1)$$

Where, (x_1, y_1) is the coordinate of one sensor node and (x_2, y_2) is the coordinate of another sensor node.

Step 3: Recalculate the centroid positions in every cluster formed and check whether there is any change in their positions compared to previous positions.

Step 4: Apply Levy flight function to optimize the clustering centroid selected by k-means to avoid to trap in local optimum.

In levy flight based clustering, a random exploration vector θ_M is added to the cluster centroid as shown in equation (2)

$$D_m^* = D_m + \theta_m \qquad (0 < m < n) \quad (2)$$

Where θ_m is the iterated random search vector in d-dimensional region, which is attained through equation (3)

$$\theta_m = rand(x_i, y_i) \times \lambda sign\left[rand - \frac{1}{2}\right] \otimes Levy, \quad (3)$$

$$i = 1,2,\ldots,D$$

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/223318 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

Where $rand()$ is a random vector and $sign\left[rand - \frac{1}{2}\right]$ signifies the direction of the exploration vector. x_i, y_i are the upper and lower limit of dimension i and λ is an influence factor that controls the search space.

Step 5: If there is any change in the position of the centroid, then go to step 2, else the chosen centroids are finalized, and the clusters are considered to be the output of this algorithm.

3.3. CH Selection

After forming the clusters, the CHs are selected for each cluster to regulate the data transmission process. The TDO algorithm is used in the proposed study to choose the best CHs for each cluster. The CH selection is relied on three main objectives: distance, energy and count of neighbors. Initially, the CH selection for the function of fitness is formulated before the process of selection.

3.3.1. Distance

To determine the sensor node closest to the target, the distance among the sensor nodes and the edge server is determined. The sensor node with a minimum distance to the edge server has the highest probability of being selected as CH. The formula to calculate the distance among the sensor node and the edge server is provided in equation (4)

$$D(N_i, N_j) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4}$$

Where, N_i *and* N_j are the sender and receiver nodes and x_i *and* y_j are the node coordinates.

3.3.2. Residual Energy

Residual energy (RE) is one of the most important parameters to consider when choosing CH to lead to energy efficiency. This value must be high for the sensor node to be selected as CH.

3.3.3. Count of Neighbors

The number of neighbors (CN) near the sensor node must be high to select the node as a CH. Therefore, this parameter must be maximized when searching for an optimal CH in the network.

3.3.4. Fitness Function

The fitness function for CH selection is computed using equation (5)

$$F = \frac{\varsigma}{D(N_i, N_j)} + \nu * RE + \upsilon * CN \tag{5}$$

Where, $\varsigma + \nu + \upsilon = 1$.

3.3.5. TDO for Optimal CH Selection

The TDO algorithm is one of the recently introduced metaheuristics inspired by the behavior of Tasmanian devils. According to this algorithm, the initial population is randomly generated in the search region, with each member of the population having an objective function. In CH selection process, the sensor nodes in the network are considered members of the TDO algorithm, and each of these sensor nodes follows an objective function. The exploration and exploitation phases of this algorithm depend on the feeding behavior of the population. In the Explore phase, the Tasmanian Devil searches for food and chooses carrion from the group. The update wording for exploration is shown in equation (6 and 7)

$$x_{i,j}^{new,E_1} = \begin{cases} x_{i,j} + r1*(c_{i,j} - \mathrm{K} * x_{i,j}); & F_{Ci} < F_i \\ x_{i,j} + r1*(x_{i,j} - c_{i,j}); & otherwise \end{cases} \tag{6}$$

$$X_i = \begin{cases} X_i^{new,E_1}; & F_i^{new,E_1} < F_i \\ X_i; & otherwise \end{cases} \tag{7}$$

Where, X_i^{new,E_1} is the new position attained by i^{th} member, $x_{i,j}^{new,E_1}$ is the new value for j^{th} variable, F_i^{new,E_1} is the value of objective function . F_{Ci} is the objective function value of designated carrion, $r1$ is a random number in [0, 1] and K is a random number which is 1 or 2.

The algorithm updates the position in the exploitation phase based on a different strategy for eating prey. The formulation can be provided as shown in equations (8 and 9)

$$x_{i,j}^{new,E_2} = \begin{cases} x_{i,j} + r1*(p_{i,j} - \mathrm{K} * x_{i,j}); & F_{Pi} < F_i \\ x_{i,j} + r1*(x_{i,j} - p_{i,j}); & otherwise \end{cases} \tag{8}$$

$$X_i = \begin{cases} X_i^{new,E_2}; & F_i^{newE_2} < F_i \\ X_i; & otherwise \end{cases} \tag{9}$$

Where, X_i^{new,E_2} is the new position of i^{th} member, $x_{i,j}^{new,E_2}$ is the value of j^{th} variable, F_i^{new,E_2} is the objective function value and F_{Pi} is the objective function value of selected prey.

All the above steps are iterated to obtain the suitable CH for each cluster. In each iteration, the proposed algorithm determines the fitness value of each node and compares it with the other nodes. The most optimal CH is selected based

RESEARCH ARTICLE

on this comparison, and from the selected CHs, the data forwarding process is performed.

3.4. Problem Formulation

After clustering the nodes in the network, the centroids in the clustering process will be ultimately selected as the CHs. The CHs in each cluster are accountable for transferring the gathered network packets from the cluster to the destination. After collecting the data from clusters, the CHs determine the appropriate path to transfer the data based on certain criteria. The proposed method evaluates the path with respect to energy consumption, execution cost and latency.

3.4.1. Energy Consumption

Energy consumption is a primary objective in routing the data packets to the endpoint. The path chosen for data transmission should utilize less energy to transfer the packets. Generally, the nodes consume energy while sending or receiving data packets in the network. The energy model utilized in [24, 25] is followed to determine the optimal path using equation (10)

$$En = \begin{cases} p\left(r_e + t_e + f_e D^2\right); & if \;\; D \le D_{th} \\ p\left(r_e + t_e + m_e D^4\right); & if \;\; D > D_{th} \end{cases} \tag{10}$$

Where, p specifies the size of the data packet, t_e and r_e specify the transmission and reception energy, f_e and m_e are the energy needed to transfer in free space and multipath, D indicates the distance and D_{th} is the distance threshold. The Euclidean distance between the CH and edge server is computed as D.

3.4.2. Execution Cost

This parameter denotes the total cost of data tranfer from CH to the edge server. When calculating execution cost, the quality of the data transmission link is taken into account. The signal to noise ratio (SNR) and distance between CH and the edge server is utilized in the computation of execution cost. The main aim of using SNR in the computation of execution cost is to enhance the overall delivery performance. When there is any failure identified in the link, it results in increased execution costs. The formulation for execution cost is taken from [26] and can be computed as shown in equation (11)

$$Ec = n_e + \left(1/D_i \right) + \left(1/SNR_i \right) \tag{11}$$

Where, n_e specifies node energy, D_i is the distance among i^{th} CH and edge server and SNR_i is the SNR of i^{th} CH. It can be computed as shown in equation (12)

$$SNR = \frac{strength\ of\ received\ signal}{background\ noise} \tag{12}$$

3.4.3. Latency

Latency indicates the delay in data transmission in the network. This can be calculated using the routing process's beginning and completion times. The latency is calculated as shown in equation (13)

$$Lt = f_t - I_t \tag{13}$$

Where, f_t specifies the ending time at which the data is transferred to its destination and I_t specifies the starting time at which the data is sensed from the lands.

3.4.4. Fitness Function

The fitness function for routing can be formulated in equation (14) as follows:

$$F = \frac{\alpha}{En} + \frac{\beta}{Ec} + \frac{\delta}{Lt} \tag{14}$$

Where, α, β and δ are constants satisfying the constraint $\alpha + \beta + \delta = 1$. The constants values are determined based on weightage of parameters such as energy, cost and latency.

3.5. Routing

The data collected by the CHs within each cluster is then forwarded to the edge servers for quality assessment. The proposed work introduces the AMGSO-based multi-objective optimization algorithm to identify the most suitable path from the available paths. This algorithm is a combination of two metaheuristics, such as the All Member Based Optimizer (AMBO) [27] and the Group Search Optimizer (GSO) [28, 29]. AMBO is a recently introduced metaheuristic algorithm for solving technical optimization problems. The developed algorithm provides more importance to each member and generates a robust routing process. It believes in the concept that even an ordinary member of the public can sometimes lead to the optimal solution. This behavior helps select the optimal path for data transmission by considering each path in the population. The GSO is one of the established algorithms inspired by the concept of animal search behavior. These two algorithms are hybridized in this work to achieve effective routing results.

3.5.1. AMGSO for Routing the Data Packets

The proposed AMGSO framework identifies the optimal path by initializing and evaluating each path available to transmit the packets. It has been found that the exploratory phase of the AMBO algorithm is more effective in exploring the search space. Also, the exploitation phase of AMBO seems to

RESEARCH ARTICLE

require a longer convergence time and sometimes leads to a locally optimal solution under complex conditions. Therefore, this phase is replaced by the exploitation behavior of the GSO algorithm to get better results. First, the paths available for transmitting the data packets are initialized as a population matrix as shown in equation (15)

$$\chi = \begin{bmatrix} \chi_1 \\ \dots \\ \chi_i \\ \dots \\ \chi_N \end{bmatrix} = \begin{bmatrix} x_1^1 & \dots & x_1^d & \dots & x_1^l \\ \dots & \dots & \dots & \dots & \dots \\ x_i^1 & \dots & x_i^d & \dots & x_i^l \\ \dots & \dots & \dots & \dots & \dots \\ x_N^1 & \dots & x_N^d & \dots & x_N^l \end{bmatrix} \tag{15}$$

Where, the population matrix is specified as χ, i^{th} member in population is specified as χ_i, x_i^d is identified as the d^{th} problem variable of i^{th} member in population, N is the total population and m specifies the count of problem parameters.

In the initial exploration phase, the AMBO algorithm is performed through a number of iterations, with each iteration evaluating the fitness function defined in Equation (16). The objective function of each path is assessed based on the values of the problem parameters. The following is an example of how to express the values that results from evaluating the objective function as a vector.

$$F(\chi) = \begin{bmatrix} F1 = F(\chi_1) \\ \dots \quad \dots. \\ Fi = F(\chi_i) \\ \dots \quad \dots. \\ FN = F(\chi_N) \end{bmatrix} \tag{16}$$

Where, $F(\chi)$ indicates the vector of objective function and Fi indicates the objective function value provided by i^{th} population member. The AMBO algorithm follows two stages to update the population: the first stage is exploration, and the second is exploitation. In the exploration phase, the algorithm updates the position values based on the position values of various routes in the problem space. The new position value is acceptable even if the new route improves the objective function value. This first stage is simulated as shown in the equations (17) to (19).

$$N_\varsigma = round\left(N \times \left(1 - \frac{t}{T}\right)\right) \tag{17}$$

Where, N_ς is the total count of selected routes leading the population, t is the present iteration and T is the total count of iterations.

$$x_i'^d = \begin{cases} x_i^d + rnd\left(x_i^d - x_k^d\right); & Fi < Fk \\ x_i^d + rnd\left(x_k^d - x_i^d\right); & else \end{cases} \quad \& k = 1 \text{ to } N_\varsigma \tag{18}$$

$$\chi_i = \begin{cases} \chi_i'; & F'i < Fi \\ \chi_i; & else \end{cases} \tag{19}$$

where, $x_i'^d$ indicates the new value obtained for the problem variable d, rnd is the random value between 0 and 1, χ_i' is the new position value attained by i^{th} route through exploration and $F'i$ is the objective function value.

Since the exploitation phase of AMBO leads to local optimal solutions, the GSO is adopted here to achieve optimal results and improve the overall convergence rate. In the exploitation phase, random migration is conducted by dispersers, and it follows systematic search strategies because of its higher efficiency. For random walks, a random head angle Φ_i is generated by the algorithm at j^{th} iteration and a random distance is chosen, which can be given in equation (20 and 21) as follows:

$$\gamma_i = c \cdot r_1 \gamma_{\max} \tag{20}$$

$$\chi_i^{j+1} = \chi_i^j + \gamma_i S_i^j \left(\Phi^{j+1}\right) \tag{21}$$

Where, γ_i is the random distance chosen, c is a constant, r_1 is a random number in the range [0, 1], χ_i^{j+1} is a new position attained, χ_i^j is the position in the previous iteration, S_i^j is the unit vector to indicate the search direction and Φ^{j+1} indicates the random head angle at iteration $j+1$. This random head angle can be computed using the following equation (22):

$$\Phi^{j+1} = \Phi^j + r\eta_{\max} \tag{22}$$

Where, r indicates a random value under uniform distribution within the range [0, 1] and $\eta_{\max}$ indicates the maximum turning angle.

The optimal path for data transfer can be determined by repeating the above procedures. In addition, the hybridization of two algorithms helps to reach the desired solution within a minimal number of iterations.

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/223318 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

3.5.1.1. Dead Node Identification

The nodes deployed in the network are random and distributed over a huge coverage area. This makes it difficult to manually identify the active and the defective sensors. The dead nodes in the network disrupt communication between the sensor nodes and edge servers, resulting in the loss of valuable information. To avoid such circumstances, the proposed work follows a constraint in which the residual energy of the nodes is used to discriminate whether the node is alive or dead. The sensors used in the network have the power to send and receive data packets. If the power is reduced then the service time of the network is also gets reduced. The algorithm finds an alternative path to send the packets when the node's residual energy reaches zero. The following equation (23) is followed by the AMGSO algorithm in each iteration to avoid communication breaks due to the presence of dead nodes.

$$n_i = \begin{cases} alive\ node;\ if\ \ R_E(n_i) > 0 \\ dead\ node;\ otherwise \end{cases} \tag{23}$$

Where, n_i indicates the i^{th} node and $R_E(n_i)$ indicates the residual energy of i^{th} node. After the identification of dead nodes, the alternative path is identified by the algorithm based on the fitness value of the routes. Using the constraint in equation (17), the proposed algorithm finds the most optimal path without any disruption in communication.

3.5.1.2. Load Balancing

Load balancing is done on the edge server side, and resource availability in the edge servers is leveraged to balance the loads across the available servers. To determine whether the respective edge server is occupied, a binary variable [30] is utilized as shown in equation (24)

$$\Psi_p = \begin{cases} 1;\ if\ \ E_p\ is\ occupied \\ 0;\ otherwise \end{cases} \tag{24}$$

Where, E_p indicates the p^{th} edge server. Another binary variable is utilized to identify whether a particular CH offloads the data to the considered edge server. The mathematical formulation for this binary variable can be given in equation (25) as follows:

$$\Omega_p^n = \begin{cases} 1;\ if\ \ CH_n\ offloads\ data\ to\ E_p \\ 0;\ otherwise \end{cases} \tag{25}$$

Where, CH_n indicates the n^{th} CH. The count of edge servers that are occupied can be identified using the equation (26)

$$N(E) = \sum_{p=1}^{P} \Omega_p^n \tag{26}$$

Based on the usage of the instances of VMs, the usage of resources in p^{th} edge server is identified, and this can be mathematically provided as shown in equation (27)

$$\Re_p = \frac{1}{\mathrm{K} \cdot w} \sum_{k=1}^{\mathrm{K}} V_p^k \tag{27}$$

Where, K indicates the rounds, w is the waiting time, V_p^k are the count of VM instances at k^{th} round in the p^{th} edge server. Thus, the average utilization of the VM instances by the edge server can be calculated using equation (28)

$$A\Re = \frac{1}{C_E} \sum_{p=1}^{P} \Re_p \tag{28}$$

Where, C_E indicates the total count of occupied edge servers in the network. The edge server's load variance is then identified to assess if the edge server is overloaded or under loaded. The edge server load variance can be determined using equation (29)

$$\mathrm{I} = \left(A\Re - \Re_p\right)^2 \tag{29}$$

The total edge server load variance can be computed as shown in equation (30)

$$A\mathrm{I} = \frac{1}{C_E} \sum_{p=1}^{P} \mathrm{I} \tag{30}$$

The edge server load is computed and evaluated for overloading and under loading issues. If an issue develops, the load is dispersed to the edge servers, and the data quality is assessed. Algorithm 1 contains the proposed AMGSO's pseudo code.

Input: Random paths

Output: Optimal path for data transmission

Initialize the algorithmic parameters, an initial population

While $(t \le T)$ do

Create an initial population based on the input paths

Evaluate the fitness function using equation (14)

Update the values of selected routes using equation (17) and update the best route based on fitness

For $(i = 1\ to\ N)$

RESEARCH ARTICLE

If $(i == dead\ \ node)$

Find an alternative path

Else

Update the position of the route using equations (18) to (19)

End if

Determine the load variance using equation (30)

End for

Store the best route identified

End while

Return optimal path

Algorithm 1 AMGSO for Optimal Path Selection

Regarding optimal path selection, the AMGSO algorithm needs to store the load balancing parameters of each route in the population, and the initial population which includes the M algorithmic parameters and the population length N. Therefore, the complexity of the AMGSO algorithm is $mathcalO(M+N+N^2)$. The algorithm may be suffer in efficiency when the sensor node density increases at massive, hence AMGSO algorithm have not much supported for scalability.

3.6. Data Quality Evaluation

After balancing the load, the QoD of the data is evaluated based on certain QoD criteria. The proposed work utilizes the measures such as completeness, correctness, precision, conciseness and reliability to determine the quality of each data provided by the CH. The completeness of the data is evaluated based on the cells and rows present in the dataset. The formulation can be given in equation (31) and (32) as follows:

$$C_1 = \left(1 - \frac{No.of\ \ incomplete\ \ cells}{No.\ of\ \ cells}\right)*100 \tag{31}$$

$$C2 = \left(1 - \frac{No.\ of\ \ incomplete\ \ rows}{No.\ of\ \ rows}\right)*100 \tag{32}$$

Both the above equations are utilized to compute the completeness of the dataset. The correctness of the incoming data is then evaluated based on the records and attributes in the dataset using the following equation (33):

$$C_R = \frac{(total\ \ records\ in\ dataset - count\ \ of\ \ errors\ identified\ )*100}{count\ \ of\ \ records\ revised * count\ \ of\ \ attributes\ \ revised} \tag{33}$$

The accuracy of the data is then calculated using the stored entries. The mathematical formula is given in equation (34) as follows:

$$P_R = \frac{count\ \ of\ \ entries\ \ stored\ \ without\ \ enough\ \ precision * 100}{count\ \ of\ \ records\ revised * count\ \ of\ \ attributes\ \ revised} \tag{34}$$

The conciseness of the incoming data is evaluated based on the duplicate values identified, and this can be mathematically modelled in equation (35) as follows:

$$C_C = \frac{(count\ \ of\ \ entries\ \ not\ used + count\ \ of\ \ entries\ \ duplicated\ )*100}{count\ \ of\ \ records\ revised * count\ \ of\ \ attributes\ \ revised} \tag{35}$$

Based on all these metrics, the overall reliability of the data is then computed using the following equation (36):

$$R_L = 100 - (100 - (C1 + C2) + 100 - C_C + \max\{100 - P_R,\ 100 - C_R\}) \tag{36}$$

Based on the reliability measure, the valid data is recognized and stored in the cloud servers, and the invalid data is discarded.

4. RESULTS AND DISCUSSION

This section provides a comprehensive illustration of the suggested framework's performance in terms of results obtained. Several experiments are conducted under diverse scenarios to prove the performance of the proposed approach. Further, the simulation scenario utilized, metrics utilized in the evaluation of the proposed methods and the analysis outcomes are detailed in the subsequent sections.

4.1. Simulation Scenario

The proposed framework follows a 3-tiered architecture that integrates the WSN, edge and cloud computing platforms for effective information gathering and communication. Initially, the WSN network is constructed for a chosen area 550×250, and the sensors are randomly deployed across the network. A total count of 250 sensors is chosen to be deployed, and each sensor is meant to sense a particular soil component.

Table 2 Parameter Settings of the Proposed Work

Parameters	Values
Network area	550×250
Count of sensors deployed	250
Components under study	pH, EC, OC, P, K, Ca, Mg, S, Zn, B, Fe, Cu and Mn
Initial energy of sensor	10J
Packets transmitted per second	343 bps
Count of edge servers deployed	3
Initial population	4816
Total number of rounds	500

The proposed work considers 13 main soil components, such as pH, EC, OC, P, K, Ca, Mg, S, Zn, B, Fe, Cu and Mn, which are recorded with suitable sensors. A total of 3 edge

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/223318 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

servers are deployed in the field far away from the sensors to collect the data from CHs and to evaluate the QoD. The valid data evaluated by the edge servers are then transferred to the cloud datacenter for storage purposes. The proposed work is evaluated using the ICRISAT publicly available dataset, and it provides the data for all the considered 13 soil components. The parameter settings followed in this work are provided in Table 2.

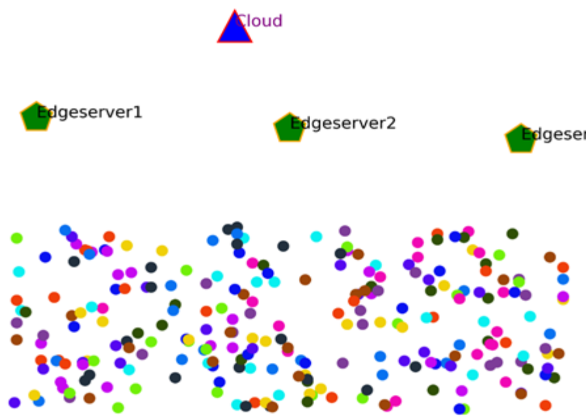


Figure 3 Sensor Deployment in the Network Region

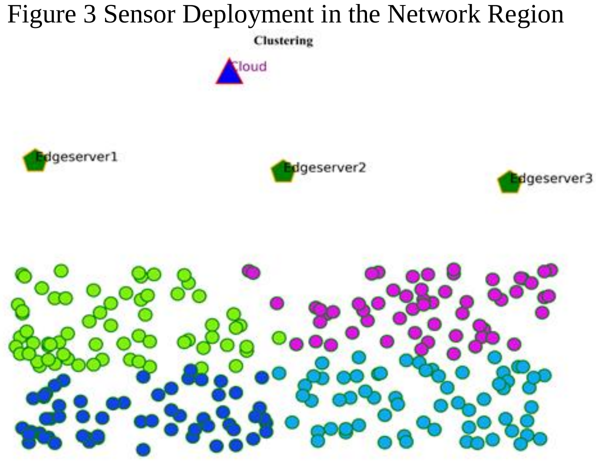


Figure 4 Clustering of Sensor Nodes in the Network

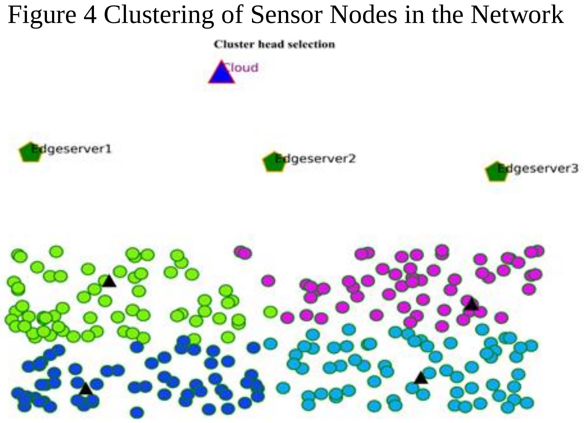


Figure 5 CH Selection Using TDO Algorithm

Figures 3-6 show the random deployment of sensors in the field, the clustering result, the CH selection process and the routing result. The entire implementations of the work are carried out in the Python platform using the Keras Tensorflow libraries. The system specification is as follows: the system considered for implementation is installed with Intel(R) Core(TM) i5-3570 processor @ 3.40 GHz. The system is installed with 8 GB RAM on a 64-bit Windows 10 pro operating system.

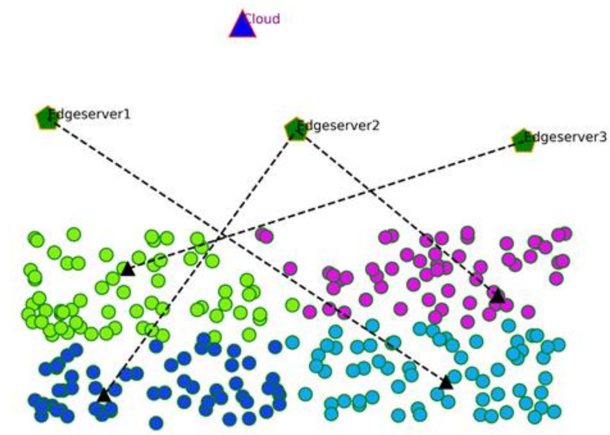


Figure 6 Routing Using the AMGSO Algorithm

4.2. Performance Metrics

The efficiency of the proposed work is evaluated in terms of energy consumption, execution cost and latency. The mathematical formulation for energy consumption is in equation (4), execution cost is in equation (5), and latency is in equation (7). Network lifetime can be computed as the time until the energy of the first node of WSN runs out. The formulation for the packet delivery ratio is provided below in equation (37):

$$PDR = \frac{Total\ \ packets\ \ received}{Total\ \ packets\ \ sent} * 100 \qquad (37)$$

Where, PDR indicates packet delivery ratio.

4.3. Performance Analysis

The overall performance analysis of the proposed work is covered in this section. The proposed work is analyzed using existing LEACH, cluster, PSQ and PHRA methods. All the results for comparison are taken from the PHRA method proposed in [16]. The performance analysis results are described below:

4.3.1. Energy Consumption

Figure 7 depicts the obtained performance of energy consumption by varying the number of nodes. The grap exhibits that the proposed strategy performs substantially superior in terms of reducing overall energy usage in the network. The graph has been plotted by varying the number of nodes between 50 and 250. The values obtained through this analysis are presented in Table 3. The proposed approach resulted in lower energy consumption in all the scenarios than in the previous works. Among the techniques compared, the PHRA technique provided optimal values, while the LEACH protocol resulted in higher energy consumption values. The optimal path identification procedure in the proposed approach has been enhanced with the addition of dead node identification and load balancing, resulting in this improvement. The overall energy consumption of the proposed approach is 0.00072J for a total of 250 nodes.

Table 3 Comparison of Energy Consumption for Proposed and Existing Methods

No of nodes	Methods				
	LEACH	Cluster	PSQ	PHRA	Proposed
50	0.00076	0.00059	0.00079	0.00014	0.00008
100	0.00111	0.00102	0.00092	0.00039	0.00030
150	0.00118	0.00110	0.00104	0.00074	0.00065
200	0.00127	0.00121	0.00117	0.00080	0.00069
250	0.00137	0.00137	0.00135	0.00087	0.00072

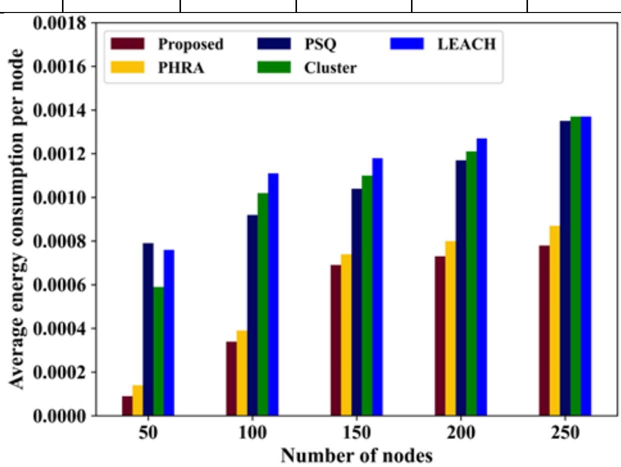


Figure 7 Energy Consumption vs Number of Nodes

4.3.2. Packet Delivery Ratio

The proposed work's packet delivery ratio is compared to current works, and the outcomes are shown in Table 4. The table shows that the values for the proposed works are higher than for the existing works. The graphical representation for this comparison is displayed in Figure 8. The figure also exhibits that the proposed approach is optimal at different node counts than the other works. Among the works compared, the PHRA technique gave an optimal packet delivery ratio, while the LEACH protocol and the cluster technique resulted in low values. Because the dead nodes in the path are accurately detected and discarded, the overall delivery ratio is improved. The packet delivery ratio of the proposed approach for 250 nodes is 97%.

Table 4 Comparison of Packet Delivery Ratio for Proposed and Existing Methods

No of nodes	Methods				
	LEACH	Cluster	PSQ	PHRA	Proposed
50	89.4	89.9	92.9	97.8	98.9
100	88.9	88.9	92.9	96.6	97.9
150	89	88.7	92	96.5	97.9
200	88.3	87.8	91.8	96	96.9
250	87.7	87.5	91.5	95.6	97

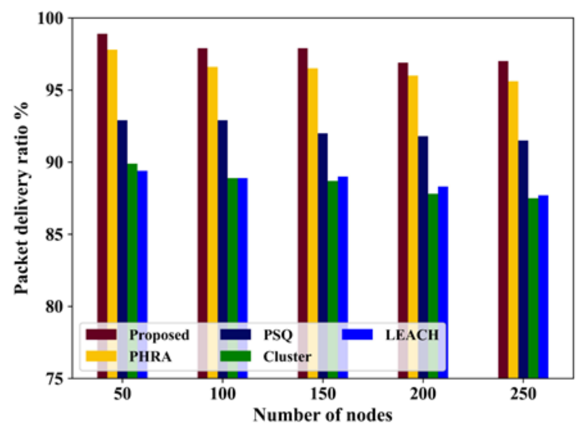


Figure 8 Packet Delivery Ratio vs Number of Nodes

4.3.3. Network Lifetime

Table 5 Comparison of Network Lifetime for Proposed and Existing Methods

No of nodes	Methods				
	LEACH	Cluster	PSQ	PHRA	Proposed
50	203.32	272.38	289	474.42	488.82
100	223.79	251.92	271.1	459.08	470.03
150	216.11	223.79	245.52	421.72	445.95
200	189.26	213.55	230.18	395.14	422.81
250	177.75	198.21	223.79	348.83	385.01

Table 5 shows the results of the network lifetime comparison. According to the table, the proposed strategy elapses the network lifetime much faster than the other comparable methods. This is further demonstrated by the graphical comparison in Figure 9. The values are plotted in the graph by varying the count of nodes between 50 and 250. The proposed approach evidently resulted in optimal values for all the

RESEARCH ARTICLE

sensor counts. The PHRA technique enhanced the network lifetime among the compared techniques, whereas the LEACH protocol resulted in lower values. The main reason for the proposed method to attain optimal values is that the dead nodes in the network are identified, and alternative paths are selected to save energy from transmissions. This concept is not considered in the compared works which resulted in this enhancement.

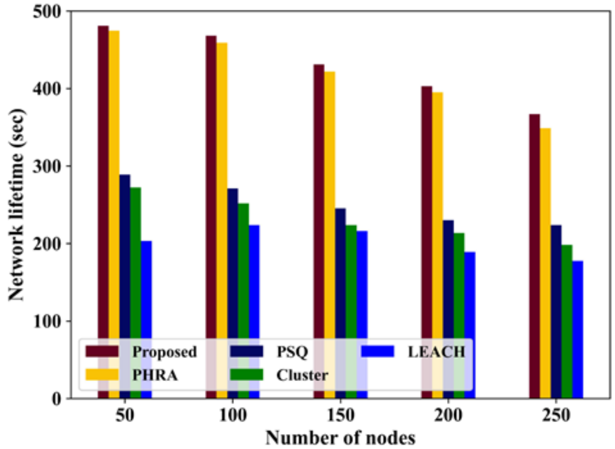


Figure 9 Network Lifetime vs Number of Nodes

4.3.4. Execution Cost

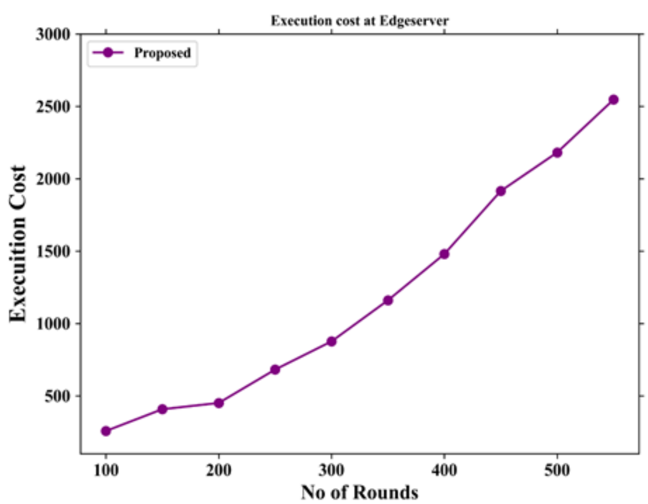


Figure 10 Execution Cost Analysis of the Proposed Method

The execution cost of the proposed methods is analyzed based on the number of rounds. The result of this analysis is graphically depicted in Figure 10. The figure exhibits an increase in execution cost when the rounds increase. The execution cost is low for 100 to 200 rounds and gradually increases when the rounds are above 200. This means that the execution cost increases for more transfers in the network.

4.3.5. Latency

The latency of the proposed method is analyzed, and the result is graphically presented in Figure 11. Also, the values obtained through this analysis are presented in Table 6. From the figure, it is observed that there is only a minor increase in latency when the nodes increase. The plot has been made for a total of 250 nodes, and when the approach reaches 250 nodes, there is a minor increase in latency. This analysis confirms that the latency of the proposed method is only minimum, even with a larger number of data transmissions. The overall simulation proves the proposed method is superior and effective than the other existing works. The proposed framework considered most of the available issues in the smart agriculture scenario to enhance the overall data acquisition process.

Table 6 Latency Comparison of the Proposed and Existing Methods

No of rounds	Methods			
	EECRP	PSO-ECHS	MEACBM	Proposed
50	0.029	0.026	0.023	0.019
100	0.03	0.027	0.024	0.019
150	0.031	0.028	0.024	0.021
200	0.031	0.028	0.025	0.021
250	0.032	0.029	0.026	0.022

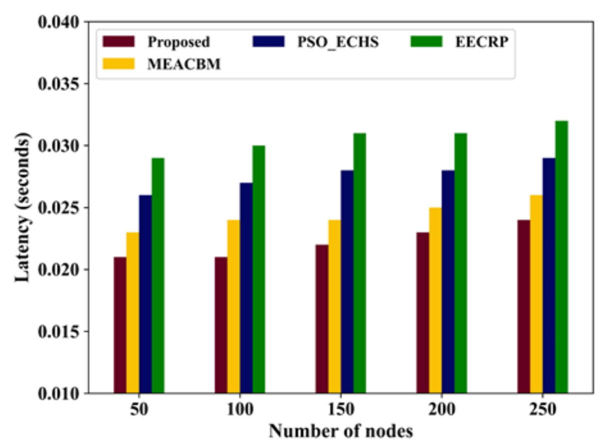


Figure 11 Latency vs Number of Nodes

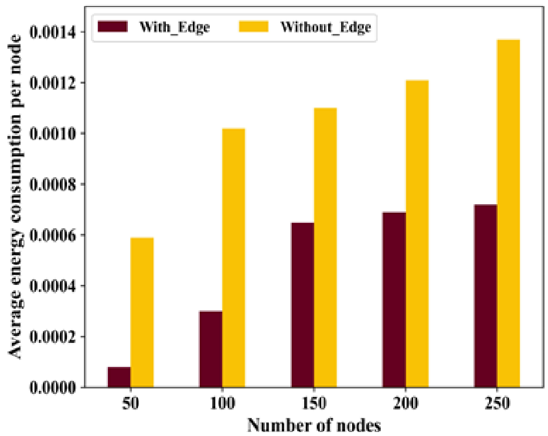
4.3.6. With and Without Edge Analysis

This section analyzes the proposed model with and without edge servers. The effect of edge computing in the proposed work is deeply analyzed, and the outcomes are interpreted. With the placement of the edge server, the data are initially forwarded to the edge server, where the QoD is evaluated. After evaluations, the data are transferred to the cloud for

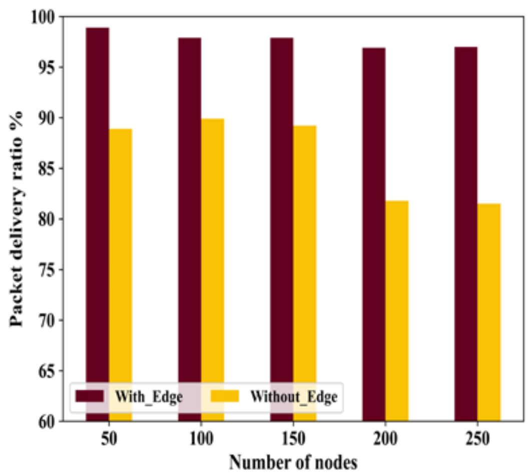
DOI: 10.22247/ijcna/2023/223318 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

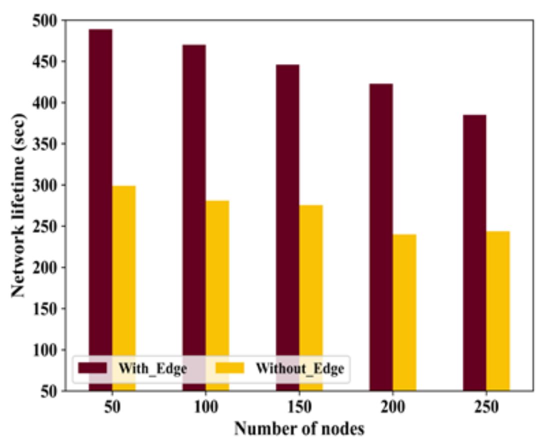
storage. In the next scenario, the edge server is removed, and the data are directly routed to the cloud server where the QoD is evaluated. After evaluations, the data are stored on the server. The evaluation results are presented below:



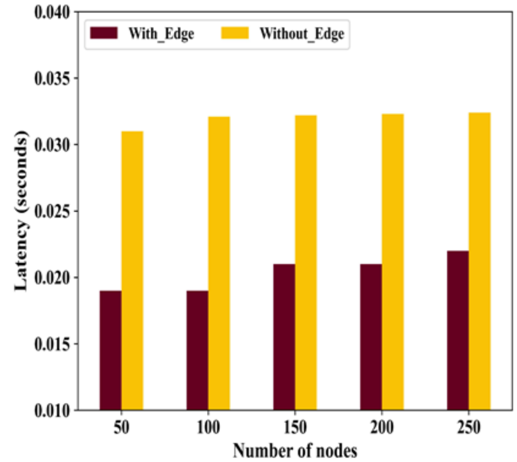
(a)



(b)



(c)



(d)

Figure 12 Performance Analysis With and Without Edge Server Based on (a) Energy Consumption, (b) Packet Delivery Ratio, (c) Network Lifetime and (d) Latency

Figure 12 depicts the results of a performance analysis of the suggested technique with and without edge server placement. The graphic shows that the placement of the edge server is critical in the proposed system since it improves overall performance. Without an edge server, the data must be sent to the cloud, which results in overhead and slows down the entire process. This causes adverse effects on the overall network lifetime and results in higher latencies for each transmission. Thus, it is clear that the combined edge and cloud computing paradigms can be utilized to achieve better performance in data acquisition.

The results indicates that most of the existing methods have more energy consumption and less network lifetime. The proposed method is energy efficient especially for agriculture precision applications. The above comparison table indicates that some systems methods moderately energy efficient but affected by network overhead. Routing protocols are occupying a significant role for data aggregation in agriculture monitoring field. Multifaceted characteristics present in the existing routing protocols demands more energy which leads to reduce the network lifetime. But the proposed routing method has been achieved better energy efficiency when compared to existing methods due to effective fitness function formulation. Moreover, the proposed routing method attained fair packet delivery rate and prolongs the lifetime of sensors.

5. CONCLUSION

This work presents an optimized data acquisition framework for smart agriculture using a hybrid optimization algorithm. The proposed work is capable of addressing the data acquisition problem in a well versed manner. Two crises that are overlooked in the existing works are resolved in this work.

RESEARCH ARTICLE

The proposed work resolved the dead node issue in the network and effectively identified the alternative path in each iteration. Moreover, the problems of overloading and under loading of data in edge servers are also resolved through load balancing. The proposed framework is implemented in Python using the ICRISAT dataset based on 13 main soil components. The evaluations proved that the proposed framework is much more efficient than the existing data acquisition models. The performance has been evaluated in terms of energy consumption, packet delivery ratio, network lifetime, latency and execution cost. The scenarios of the evaluations are varied by changing the number of nodes and the total number of rounds used. Under each scenario, the proposed algorithm resulted in effective outcomes compared to other works.

In future, it is planned to conduct more experiments with real time datasets and big data to prove the practicality of the proposed framework. Also, a single CH is elected for each cluster in this work which may affect the communication as the CH may die when the energy is depleted. Thus, CH rotation can be considered after CH selection, where the CHs are elected in every round based on the residual energy maintained. Apart from these, this work considers single-hop communication, which restricts the extension of WSN for larger fields. Therefore, it is also recommended to consider multi-hop communication in the future for larger areas.

REFERENCES

[1] S. K. Dhillon, C. Madhu, D. Kaur, S. Singh, "A review on precision agriculture using wireless sensor networks incorporating energy forecast techniques." Wireless Personal Communications vol. 113, no, 4, pp. 2569-2585, 2020.

[2] C. Venkataramanan, S. Ramalingam, A. Manikandan, "LWBA: Levy-walk bat algorithm based data prediction for precision agriculture in wireless sensor networks." Journal of Intelligent & Fuzzy Systems Preprint pp. 1-14, 2021.

[3] A. Deroussi, I. Alihamidi, L. A. Charaf, A. A. Madi, A. Addaim, "Routing Protocols for WSN: A Survey Precision Agriculture Case Study." InIEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS) pp.1-7, 2020.

[4] V. G. Chavan, P. R. Patil, A. D. Potgantwar, P. R. Bhaladhare, "A Novel Intelligent Efficient and Dynamic Cluster Based Routing Protocol for IoT Assisted Precision Agriculture." Journal Of Algebraic Statistics vol. 13, no. 3, pp. 4025-4037, 2022.

[5] J. Lloret, S. Sendra, L. Garcia, J. M. Jimenez, "A Wireless Sensor Network Deployment for Soil Moisture Monitoring in Precision Agriculture." Sensors vol. 21, no. 21, pp. 7243, 2021.

[6] A. Chehri, H. Chaibi, R. Saadane, N. Hakem, M. Wahbi, "A framework of optimizing the deployment of IoT for precision agriculture industry." Procedia Computer Science vol. 176, pp. 2414-2422, 2020.

[7] R. K. Singh, R. Berkvens, M. Weyn, "AgriFusion: An architecture for IoT and emerging technologies based on a precision agriculture survey." IEEE Access vol. 9, pp. 136253-136283, 2021.

[8] N. N. Malik, W. Alosaimi, M. I. Uddin, B. Alouffi, H. Alyami, "Wireless sensor network applications in healthcare and precision agriculture." Journal of Healthcare Engineering 2020, 8836613, 2020.

[9] B. N. Alhasnawi, B. H. Jasim, B. A. Issa, "Internet of things (IoT) for smart precision agriculture." IJEEE vol. 16, pp. 28-38, 2020.

[10] D. K.Anguraj, V. N. Mandhala, D. Bhattacharyya, T. Kim, "Hybrid neural network classification for irrigation control in WSN based precision agriculture." Journal of Ambient Intelligence and Humanized Computing pp. 1-12, 2021.

[11] J. Alejandrino, R. Concepcion, V. J. Almero, M. G. Palconit, A. Bandala, E. Dadios, "A hybrid data acquisition model using artificial intelligence and IoT messaging protocol for precision farming." In IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) pp. 1-6, 2020.

[12] A. Jana, A. Roy, "Remotely Data Acquisition In Precision Agriculture: A Secure Way Of Data Transmission." Int. Journal of Sci & Tech Research vol.9,no.3,pp.874-878,2020.

[13] H. K. D. Sarma, "Data Aggregation in Internet of Things aiming at Precision Agriculture." In IEEE International Symposium on Smart Electronic Systems (iSES)(Formerly iNiS) pp. 271-276, 2021.

[14] K. Ghosh, S. Sharma, 'Fermat Point-Based Wireless Sensor Networks: A Default Choice for Measuring and Reporting Farm Parameters in Precision Agriculture." In Smart Agriculture Automation Using Advanced Technologies pp. 141-150, 2021. Springer, Singapore,

[15] A. C. Tagarakis, D. Kateris, R. Berruto, D. Bochtis, "Low-cost wireless sensing system for precision agriculture applications in orchards." Applied Sciences vol. 11, no. 13, pp. 5858, 2021.

[16] F. Khelifi, "Monitoring system based in wireless sensor network for precision agriculture." In Internet of Things (IoT) pp. 461-472, 2020. Springer, Cham,

[17] C. Li, D. Chen, C. Xie, Y. Tang, "Algorithm for wireless sensor networks in ginseng field in precision agriculture." Plos one vol. 17, no. 2, pp. e0263401, 2022.

[18] J. Agarkhed, P. Y. Dattatraya, S. Patil, "Precision agriculture with cluster-based optimal routing in wireless sensor network." International Journal of Communication Systems vol. 34, no. 10, pp. e4800, 2021.

[19] V. Pandiyaraju, R. Logambigai, S. Ganapathy, A. Kannan, "An energy efficient routing algorithm for WSNs using intelligent fuzzy rules in precision agriculture." Wireless Personal Communications vol. 112, no. 1, pp. 243-259, 2020.

[20] K. N. Qureshi, M. U. Bashir, J. Lloret, A. Leon, "Optimized cluster-based dynamic energy-aware routing protocol for wireless sensor networks in agriculture precision." Journal of sensors 2020.

[21] Y.D. Yao, X. Li, Y.P. Cui, J.J. Wang, C. Wang, "Energy-efficient routing protocol based on multi-threshold segmentation in wireless sensors networks for precision agriculture." IEEE Sensors Journal vol. 22, no.7, pp. 6216-6231, 2022.

[22] H.B. Mahajan, A.A. Junnarkar, M. Tiwari, T. Tiwari, M. Upadhyaya, "LCIPA: Lightweight clustering protocol for industry 4.0 enabled precision agriculture." Microprocessors and Microsystems vol. 94, pp. 104633, 2022.

[23] A. Kumar, B.S. Dhaliwal, D. Singh, "CL-HPWSR: Cross-layer-based energy efficient cluster head selection using hybrid particle swarm wild horse optimizer and stable routing in IoT-enabled smart farming applications." Transactions on Emerging Telecommunications Technologies vol. 34, no. 3, pp. e4725, 2023.

[24] K. M. Awan, A. Ali, F. Aadil, K. N. Qureshi, "Energy efficient cluster based routing algorithm for wireless sensors networks." In IEEE International Conference on Advancements in Computational Sciences (ICACS) 1-7, 2018.

[25] J. Shen, A. Wang, C. Wang, P. CK Hung, C.F. Lai, "An efficient centroid-based routing protocol for energy management in WSN-assisted IoT." Ieee Access vol. 5, pp. 18469-18479, 2017

[26] K. Haseeb, I. Ud Din, A. Almogren, N. Islam, "An energy efficient and secure IoT-based WSN framework: An application to smart agriculture." Sensors vol. 20, no. 7, pp. 2081, 2020.

[27] F. A. Zeidabadi, S. A. Doumari, M. Dehghani, Z. Montazeri, P. Trojovsky, G. Dhiman, "AMBO: all members-based optimizer for solving optimization problems." CMC-Comput. Mater. Contin vol. 70, pp. 2905-2921, 2022.

[28] S. He, Q. H. Wu, J. R. Saunders, "Group search optimizer: an optimization algorithm inspired by animal searching behavior." IEEE

RESEARCH ARTICLE

transactions on evolutionary computation vol. 13, no. 5, pp. 973-990, 2009.

[29] L. Abualigah, "Group search optimizer: a nature-inspired meta-heuristic optimization algorithm with its results, variants, and applications." Neural Computing and Applications vol. 33, no. 7, pp. 2949-2972, 2021.

[30] J. Zhang, X. Li, X. Zhang, Y. Xue, G. Srivastava, W. Dou, "Service offloading oriented edge server placement in smart farming." Software: Practice and Experience vol. 51, no. 12, pp. 2540-2557, 2021.

Authors

Vuppala Sukanya received the M.Tech. Degree in computer science and engineering from the IIIT Banagalore in 2007. She is currently pursuing Ph.D. from Osmania University. She has been an Assistant Professor with the Department of Computer Science and Engineering, Osmania University since 2013. Her research interests include wireless sensor networks, Internet of Things and Cloud Computing. Email: sukanya.v@uceou.edu.

Ramachandram S is currently working as Vice Chancellor of Anurag University. He completed Ph.D. (Processing of Read-Only Transactions in Mobile Broadcast Environment), 2005, from Osmania University, Hyderabad. His research areas include Distributed and Cloud Computing, Cyber Security. He has authored books on Software Engineering, Operating Systems. He has published about 110 research papers in various journals and conferences. He is a senior member of Institute of Electrical and Electronic Engineers (IEEE), Computer Society of India (CSI), ACM. He was elected a fellow of Institute of Electronics and Telecommunication Engineers (IETE) and a fellow of Telangana Academy of Sciences. He was Awarded 'Eminent Engineering Teacher' by Institution of Engineers (India), 2016. Email: schandram@gmail.com.

How to cite this article:

Vuppala Sukanya, Ramachandram S, "A Novel All Members Group Search Optimization Based Data Acquisition in Cloud Assisted Wireless Sensor Network for Smart Farming", International Journal of Computer Networks and Applications (IJCNA), 10(4), PP: 651-667, 2023, DOI: 10.22247/ijcna/2023/223318.